

(19)



Europäisches Patentamt

European Patent Office

Office européen des brevets



(11)

**EP 0 794 250 A1**

(12)

**EUROPEAN PATENT APPLICATION**

(43) Date of publication:

**10.09.1997 Bulletin 1997/37**

(51) Int. Cl.<sup>6</sup>: **C12N 15/53**, C12N 15/82,

C12N 9/02, C12N 5/10,

C12Q 1/68

// A01H5/00

(21) Application number: **97103098.6**

(22) Date of filing: **26.02.1997**

(84) Designated Contracting States:

**BE ES FR GB IT NL**

(30) Priority: **04.03.1996 CH 550/96**

(71) Applicants:

• **SOREMARTEC S.A.**

**6700 Arlon-Schoppach (BE)**

Designated Contracting States:

**BE ES FR GB NL**

• **Ferrero S.p.A.**

**I-12051 Alba (Cuneo) (IT)**

Designated Contracting States:

**IT**

(72) Inventors:

• **Dani, Maria**

**81020 Sala di Caserta (IT)**

• **Catello, Sergio**

**80129 Napoli (IT)**

(74) Representative: **Rambelli, Paolo et al**

**c/o JACOBACCI & PERANI S.p.A.**

**Corso Regio Parco, 27**

**10152 Torino (IT)**

(54) **Isolation and sequencing of the hazel FAd2-N gene**

(57) The invention relates to the isolation from hazel (*Corylus avellana* L.) of the FAD2-N gene coding for the  $\Delta 12$  desaturase enzyme of the microsomal fraction and, in particular, provides the nucleotide sequence and the deduced amino-acid sequence of the gene and provides for its use as a probe for the isolation of other plant desaturases. It also relates to the use of this gene for altering the desaturase levels and consequently the fatty-acid composition of the plant.

**EP 0 794 250 A1**

## Description

The present invention relates to the isolation from hazel (*Corylus avellana* L.) of the FAD2-N gene which codes for the  $\Delta 12$  desaturase enzyme of the microsomal fraction.

More particularly, the invention relates to the nucleotide sequence, to the derived amino-acid sequence of the gene, and to its use as a probe for the isolation of other plant desaturases. It also relates to the use of this gene for altering the desaturase levels, and consequently the fatty-acid composition of the plant.

Alteration of the fatty-acid composition may have various applications in the industrial field. One of the greatest problems with hazelnuts is that they become rancid by oxidation. This is due to the auto-oxidation of unsaturated lipids with the consequent formation of volatile substances with a rancid odour which cannot easily be eliminated by the usual preservation systems. Amongst the possible strategies for reducing the tendency to become rancid, the best seems to be that of reducing the degree of unsaturation of the fatty acids present in the kernel oil, since susceptibility to auto-oxidation is positively correlated with this parameter. In fact, the rate of peroxide formation is correlated with the number of C=C double bonds in the fatty acids. The rate of auto-oxidation of the fatty acids in comparison with the oleate (18:1) is about 30 times greater in the linoleate (18:2) and 80 times greater in the linolenate (18:3). Moreover, the volatile substances resulting from the degradation of the linoleate and of the linolenate have a lower threshold of perception than those derived from the oleate. A reduction in linoleic acid should reduce the availability of substrates for lipoxygenase, reduce the loss of vitamin E during preservation, and reduce the production of volatile substances such as hexanals.

In the angiosperms, most of the synthesis of polyunsaturated lipids takes place by means of a single enzyme, that is,  $\Delta 12$  (or  $\omega 6$ ) desaturase (18:1 desaturase), of the endoplasmic reticulum, although there is an 18:1 chloroplast desaturase in the leaves of some plants. Moreover, this enzyme is responsible for more than 90% of the synthesis of polyunsaturated fatty acids in non-photosynthetic tissues such as, for example, in the kernels. The conversion of oleic acid (18:1) to linoleic acid (18:2) thus takes place by means of  $\Delta 12$  desaturase, and from linoleic acid to linolenic acid (18:3) by means of  $\Delta 15$  (or  $\omega 3$ ) desaturase.

It has been shown with mutants of *Arabidopsis* that the FAD2 locus contains a gene which codes for the oleate desaturase enzyme of the endoplasmic reticulum (Okuley et al, 1994, The Plant Cell 6, 147-158). The FAD2 gene was in fact able to complement mutants of *Acabidopsis* which were deficient in desaturase activity of the endoplasmic reticulum. The gene coding for the same enzyme in soya has also recently been isolated and sequenced (Heppard et al, 1995, Plant Physiol., in press).

A reduction in the  $\Delta 12$  desaturase levels should therefore lead to a reduction in the linoleic acid content and, as a secondary effect, probably also to a reduction in linolenic acid. In hazelnuts the percentage of linoleic acid varies from 5 to 15%; the percentage of linolenic acid is from 0.1 to 0.2%. A reduction in these fatty acids should therefore be useful in the preservation of hazelnuts. There is therefore clearly a need to isolate the gene which codes for the  $\Delta 12$  desaturase of the endoplasmic reticulum. The sequence of the gene could thus be used for gene inactivation in hazelnut kernels. This inactivation could be carried out either by the antisense technique (Smith et al. (1988) Nature 334, 724-726) or by the "transwitch" technique (Flavell (1994) Proc. Natl. Acad. Sci. USA 91, 3490-3496). In the antisense technique, the hazel would have to be transformed by the entire FAD2-N gene or by portions thereof, inserted in the opposite direction to the regulating sequences. In the "transwitch" technique, the hazel would have to be transformed by an identical copy of the FAD2-N gene.

The subjects of the present invention are defined by the following claims.

Embodiments of the present invention will now be described with reference to the following drawings, in which:

Figure 1 shows the restriction map of the N2 genome clone,

Figure 2 shows the nucleotide sequence of the hazel FAD2-N gene; the amino-acid sequence of the coding portion is also shown;

Figure 3 shows the nucleotide sequence of the "I" clone of cDNA,

Figure 4 shows a comparison between the nucleotide sequences of the "I" and "N2" clones,

Figure 5 shows a comparison between the amino-acids of the "N2" gene and  $\Delta 12$  desaturases of *Arabidopsis* and of soya,

Figure 6 shows the homology between hazel  $\Delta 12$  desaturase and various desaturases of other plants both plastid and of the endoplasmic reticulum,

Figure 7 shows the expression of the N2 gene in various varieties of hazel both in the leaves and in the kernels.

Isolation and cloning of the FAD2 gene of *Arabidopsis thaliana* for use as a probe

In order to isolate the gene which codes for hazel  $\Delta 12$  desaturase enzyme, it was necessary to use the FAD2 gene of *Arabidopsis* as a probe.

In order to isolate the *Arabidopsis* gene, two oligonucleotides were used as "primers" for the amplification of the sequences included between the start and the end of the gene. The oligonucleotides used were NOCC1 (CTGAATTC-CAGGTGGAAGAATGCC) which contains the Eco RI restriction site and the sequences corresponding to the portion between bases 100 and 116 of the gene (Okuley J. et al, 1994, The Plant Cell 6, 147-158) and NOCC4 (AGGAATTC-GACAATTTCTTCACCATCATGC) which contains the restriction site of the Eco RI enzyme and the sequences complementary to the portion between base 1245 and base 1266. The amplification reaction was as follows: 12.8  $\mu$ l H<sub>2</sub>O, 2.5  $\mu$ l 10 x PCR buffer (Perkin Elmer), 2.5  $\mu$ l *Arabidopsis* genome DNA (10 ng/l), 1  $\mu$ l dNTP, each 2.5mM, 2  $\mu$ l 25mM MgCl<sub>2</sub>, 1  $\mu$ l NOCC1 oligonucleotide (50ng/ $\mu$ l), 1  $\mu$ l NOCC4 oligonucleotide (50ng/ $\mu$ l) 0.2  $\mu$ l Taq I DNA polymerase (Perkin Elmer) (5U/ $\mu$ l). The mixture thus prepared was subjected to 1 denaturing cycle for 1 minute at 94°C and to 40 cycles composed as follows: 30 seconds at 94°C, 1 minute at 52°C, 2 minutes at 72°C. The amplification products were separated on 1% agarose gel in TAE buffer (0.04M Tris-acetate, 0.002M EDTA) and stained with ethidium bromide at a concentration of 0.5  $\mu$ g/ml. The portion of gel containing the fragment of the expected length was withdrawn. In order to extract the DNA, 10  $\mu$ l of Qiaex resin (Qiaex extraction kit, firm Qiagen) were added for each 200mg of gel. The supplier's method was then followed. The DNA was then supplemented with a tenth of a volume of 10XH buffer (Boehringer) and 20 units of Eco RI enzyme (Boehringer). After incubation overnight at 37°C, the DNA was precipitated with 0.1 volumes of 5M NH<sub>4</sub>OAc and one volume of isopropanol. After 10 minutes at ambient temperature, the DNA was centrifuged for 20 minutes at 14000 rpm and the precipitate was washed with 70% ethanol. The DNA was resuspended in 15  $\mu$ l of H<sub>2</sub>O. The concentration was determined on gel by comparison with a known standard.

The amplified fragment was inserted in the pUC18 vector. A ligation mixture was prepared as follows: 1  $\mu$ l pUC18 plasmid DNA cut with Eco RI (20ng), 1.5  $\mu$ l fragment amplified with NOCC1 and 4 (25ng), 1  $\mu$ l 10X ligase buffer (Boehringer), 1  $\mu$ l T4 DNA ligase (1U/ $\mu$ l) (Boehringer), 4.5  $\mu$ l H<sub>2</sub>O. The reaction mixture was incubated at 14°C for 12 hours.

In order to prepare competent cells, the method based on the compound hexamino-cobalt chloride was used (Maniatis, Molecular cloning, 1989, Cold Spring Harbor Laboratory Press, 1.76-1.81). 10  $\mu$ l of the ligation mixture were added to each aliquot of competent cells, defrosted on ice. After the cells had been incubated on ice for 30 minutes they were subjected to thermal shock at 42°C for 90 seconds and were then replaced in ice for 60 seconds. After the addition of 0.5 ml of SOC broth (2% Bactotryptone, 0.5% yeast extract, 10mM NaCl, 2.5mM KCl, 10mM MgCl<sub>2</sub>, 20mM glucose, pH7), the cells were incubated at 37°C with stirring for 90'. 100, 200 and 300  $\mu$ l aliquots were spread on plates containing solid LB broth (10gr/l NaCl, 10gr/l Bactotryptone, 5gr/l yeast extract, pH7.5, 15gr/l agar) with the addition of 50  $\mu$ g/ml of ampicillin and in the presence of IPTG and X-Gal. The plates were then incubated at 37°C overnight.

Some of the bacterial colonies obtained were first analyzed for their plasmid content by a quick method (Maniatis, Molecular cloning, 1989, Cold Spring Harbor Laboratory Press, 1.32). The colonies containing a plasmid of the expected length were grown and their plasmid DNA extracted (Maniatis, Molecular cloning, 1989, Cold Spring Harbor Laboratory Press, 1.33). Those containing a fragment of the expected length (1160 bp) were identified by digestion of the plasmid DNA with Eco RI. The E1 colony was selected.

One end of the insert of the E1 colony was sequenced. The plasmid DNA of the E1 clone was denatured and partially sequenced by Sanger's method using the enzyme Sequenase and <sup>35</sup>S-dATP (Amersham). The sequencing products were separated on 8% acrylamide, 8M urea, 1XTBE gel. After electrophoresis, the gel was dried and exposed overnight in contact with an autoradiographic plate ( $\beta$  max, Amersham). The sequence was compared with that published and was identical, identifying the *Arabidopsis* FAD2 gene in the cloned fragment.

#### Extraction of nucleic acids from hazel

Hazelnuts of the Nocchione, Montebello and San Giovanni varieties were harvested when almost fully ripe. The kernel was skinned before being used or frozen in liquid nitrogen. The leaves were harvested at a young stage and frozen in liquid nitrogen. 3 ml of extraction buffer were used for each gram of vegetable material with the use of the method described by Verwoerd et al. (Nucl. Ac. Res., 1989, 2362). Upon completion of the extraction, two selective precipitations were carried out by the addition of NaCl 2M, and 2 volumes of 95% ethanol to eliminate polysaccharides. The final pellet was resuspended in H<sub>2</sub>O. Further centrifuging was then carried out to eliminate any non-resuspended material.

On the other hand, DNA was extracted from young leaves of the Nocchione and Montebello varieties. The vegetable tissue was pulverized in liquid nitrogen and the DNA extracted by the CTAB (REF) method. To eliminate the polysaccharides, NaCl 2M and 2 volumes of 95% ethanol were added. The samples were incubated for 15' at -80°C and centrifuged for 15' at 4°C and 14000 rpm (Eppendorf). This selective precipitation was repeated twice and the final pellet was resuspended in H<sub>2</sub>O. Further centrifuging was then carried out to eliminate any non-resuspended material.

Checking of the probe on hazel DNA and RNA

About 20µg of DNA of the Montebello and Nocchione varieties was cut with Eco RI restriction enzyme in a volume of 300 µl in the presence of 400 units of enzyme and H buffer (Promega), with incubation for one night at 37°C. After digestion had been checked by gel electrophoresis of one twentieth of the reaction mixture, the samples were precipitated with ethanol and resuspended in 30µl of H<sub>2</sub>O. The DNA was then subjected to electrophoresis on 0.7% agarose gel and transferred by capillarity onto nylon membrane (Southern blot) for one night in the presence of 20 x SSC (3M NaCl, 0.3M Na citrate). The membrane was dried in air for 30' and then fixed by UV treatment (120,000 µJ/cm<sup>2</sup>).

The *Arabidopsis* Δ12 desaturase gene was used as a probe. For this purpose, the plasmid DNA of the E1 clone (5µg) was cut with 20 units of Eco RI in the presence of H buffer (Boehringer) in a volume of 30µl for 12 hours at 37°C. The insert of the clone was separated from the vector by electrophoresis on 1% agarose gel and extracted from the gel with the use of Qiaex resin in accordance with the suppliers' instructions (Qiagen). The DNA was denatured for 10' at 100°C, cooled rapidly in dry ice, and marked by the random priming method with the use of 6000 Ci/mmol (α<sup>32</sup>)P dATP and the reagents of Boehringer's marking kit.

The nylon membrane containing the hazel DNA was prehybridized for 1.5 hours at 55°C in standard buffer (5 x SSC, 0.1% (w/v) N-laurylsarcosine, 0.02% SDS, 1% blocking reagent solution) (10% blocking reagent solution: 10gr Boehringer blocking reagent in 150mM NaCl, 100mM maleic acid, pH7.5). The membrane was then hybridized with the *Arabidopsis* probe for one night at 55°C. The non-hybridized probe was washed twice for 15' in 2 x SSC, 0.1% SDS and twice for 15' each in 0.3 x SSC, 0.1% SDS, always at a temperature of 55°C. The probe remained coupled to the homologous sequences on the membrane was detected by autoradiography.

The RNA extracted from the young leaves of the Montebello and Nocchione varieties and from the kernels of the San Giovanni variety was separated on denaturing gel in the presence of formamide and transferred to nylon membrane by Northern blotting (Maniatis, Molecular cloning, 1989, Cold Spring Harbor Laboratory Press, 7.43-7.45). 40µg/sample of total RNA extracted from San Giovanni kernels, Nocchione leaves and Montebello leaves were used. 60 pg of probe were used as a positive control. The RNA was loaded onto a 1% agarose gel in the presence of formaldehyde. The samples were then subjected to electrophoresis for 3 hours at 80 volts in the presence of 1xMOPS. The gel was rinsed in H<sub>2</sub>O and then stained with ethidium bromide 0.5 µg/ml to display the RNA. The RNA was then transferred onto a nylon membrane (Boehringer) by "capillary blotting" in the presence of 20 x SSC throughout the night at 4°C. After transfer, the membrane was dried on 3 MM paper and then fixed by crosslinking using UV light (Stratagene UV Stratalinker 120000 µJ/cm<sup>2</sup>). The RNA was hybridised with the *Arabidopsis* Δ12 desaturase probe as described for the DNA. Detection was carried out by autoradiography. The heterologous *Arabidopsis* probe was able to display a band with a molecular weight of about 1500 bp in the hazel RNA and 3 bands of about 18, 8 and 2.8 kb in the hazel DNA cut with Eco RI.

Construction of a gene library of cDNA

The gene library of cDNA was constructed from RNA from kernels harvested when almost fully ripe and taken from plants of the San Giovanni variety. For this purpose, the Poly(A)+mRNA was isolated from the total RNA with the use of the Poly(A)Tract mRNA Isolation System II, in accordance with the method provided by the firm Promega. The samples were eluted in H<sub>2</sub>O and precipitated with 0.1 volumes of 3M NaOAc and 3 volumes of 95% ethanol. After one night at -80°C, the RNA was centrifuged for 15' at 14000 rpm (Eppendorf), the pellet was rinsed in 75% ethanol and resuspended in 10µl of H<sub>2</sub>O. The concentration was read with a spectrophotometer and the yield was 3.2µg of Poly(A)+mRNA per mg of total RNA.

The messenger RNA polyadenilate derived from kernels of the San Giovanni variety was used as a template for the synthesis of complementary DNA (cDNA) with the use of Boehringer's "cDNA synthesis kit" in accordance with the method recommended by the suppliers. An extraction was then carried out with one volume of phenol:chloroform: isoamyl alcohol (25:24:1). The cDNA was then purified in a Pharmacia column (cDNA spun columns) after the addition of NaCl 100 mM. The buffer used was the following: 10mM Tris-HCl pH 7.5, 1mM EDTA, 150mM NaCl. Eco RI "adaptors" (Pharmacia) were added to the ends of the cDNA. The reaction mixture contained: 5µl of cDNA (half of the cDNA obtained from 6µg of Poly(A)+RNA), 10µl of ligase buffer 10 x (Promega), 10µl of Eco RI adaptors (0.01u/µl), 6 units of T4 DNA ligase (Promega), in a final volume of 100µl. After incubation for 12 hours at 12°C, the ligase enzyme was inactivated for 10' at 65°C. Phosphorylation of the adaptors then followed by the addition, to the 100µl mixture, of 10µl of 100mM ATP and 10 units of T4 polynucleotide kinase. After incubation at 37°C for 30', the enzyme was inactivated by incubation for 10' at 65°C. Purification was then carried out with one volume of phenol:chloroform:isoamyl alcohol (25:24:1). The cDNA was then purified from fragments of less than 400 bp as follows. After the addition of NaCl to a final concentration of 0.1M NaCl, the cDNA was separated by chromatography in a column with Sepharose CL-4B resin (Size prep 400 spun column, Pharmacia) according to the method suggested by the suppliers. The fragments of cDNA shorter than 400 bases were thus excluded. The cDNA was precipitated with one thirtieth of a volume of 3M NaOAc and 2 volumes of 95% ethanol, centrifuged and resuspended in 10µl of H<sub>2</sub>O.

The cDNA was inserted in the  $\lambda$  phage vector Zap II cut with Eco RI and dephosphorylated (Stratagene) in the following manner: 2  $\mu$ l of cDNA (200 ng), 1  $\mu$ l of  $\lambda$  Zap II cut with Eco RI (1  $\mu$ g/ $\mu$ l) (Stratagene), 0.5  $\mu$ l of T4 DNA Ligase (4U/ $\mu$ l) (Promega), 0.5  $\mu$ l of 10 x ligation buffer (Promega), 1  $\mu$ l of H<sub>2</sub>O. The reaction mixture was incubated for 14 hours at 12°C. The mixture containing the cDNA inserted in the vector was used for the reconstruction of the phages with the use of Stratagene's Gigapack Gold "in vitro packaging" kit. The gene library of phages thus obtained was constituted by about 300,000 pfu (plaque-forming units). In order to amplify the gene library, XL1 Blue MRF' cells were prepared as described by Stratagene and used the same day. The gene libraries were plated at a concentration of about 5000 pfu per plate (95 cm<sup>2</sup>). After growth, the phages were resuspended in SM (5.8gr/l NaCl, 2gr/l MgSO<sub>4</sub>·7H<sub>2</sub>O, 50ml/l 1M Tris HCl (pH 7.5), 5ml/l 2% gelatine) and, after the addition of chloroform to 5% and incubation for 15 minutes at ambient temperature, the cell debris was centrifuged for 10 minutes at 2000 x g. Chloroform to 0.3% was added to the supernatant liquid and the phages were preserved at 4°C. Aliquots were preserved at -70°C after the addition of DMSO to 7%. The gene library was titled.

#### Construction of a partial genome gene library

The DNA of the Nocchione variety was digested with Eco RI restriction enzyme and separated on agarose gel. The fragments with lengths of up to 10000 bp (base pairs) were isolated from the gel with the use of Qiaex resin according to the Qiagen's method. For cloning in the  $\lambda$  vector Zap II, 400ng of DNA fragments were incubated with 1  $\mu$ g of desphosphorylated  $\lambda$  Zap II (Stratagene) in the presence of ligase buffer and 1.5 units of T4 DNA ligase (Promega) for 12 hours at 14°C.

Stratagene's Gigapack Gold "in vitro packaging" kit was used in accordance with the suppliers' instructions to make up the gene library. The gene library of phages thus produced was amplified as described for the cDNA gene library. The complexity of the gene library was 1,500,000 clones. This gene library was also amplified.

#### Screening of the cDNA gene library

About 250,000 phages of the cDNA gene library were plated on LB broth in the presence of XL1 Blue MRF' cells, divided into 12 plates each containing 20,000 pfu. After growth, the phages were transferred onto nylon membranes and their denatured DNA was fixed on the membranes as described by Boehringer for screening with non-radioactive probes. The membranes were then hybridized with the *Arabidopsis*  $\Delta$ 12 desaturase gene. The probe was prepared by the isolation of the insert containing the entire coding region of the gene from the plasmid. The insert was then marked with digoxigenin-dUTP with the use of Boehringer's "DNA labelling kit". Prehybridization was carried out in standard buffer (Boehringer) and hybridization was carried out in the same buffer with the addition of the *Arabidopsis* probe at a concentration of 10ng/ml and at a temperature of 55°C.

After washing twice in 2xSSC, 0.1% SDS for 5 minutes at ambient temperature and washing twice in 0.3xSSC, 0.1%SDS at 55°C, detection was carried out with the use of an anti-digoxigenin antibody conjugated with alkaline phosphatase (Boehringer) and a chemiluminescent substrate (AMPPD, Boehringer).

11 positive phage plaques were identified. These were isolated, the phages resuspended in SM and titled. From 50 to 200 phages were plated for each positive plaque. The plaques were transferred onto nylon membranes and subjected to a second hybridization with the *Arabidopsis*  $\Delta$ 12 desaturase probe, as already described above. The following clones which could hybridize with the *Acabidopsis*  $\Delta$ 12 desaturase gene were obtained from the second screening: I, F, 4.

#### Screening of the genome gene library

The gene library of Nocchione DNA was subjected to screening in the same way as the cDNA gene library. 1,600,000 phages were plated, divided into 40 plates. After growth, they were transferred to nylon membranes as described for the cDNA gene library. The membranes were then hybridized with the *Arabidopsis*  $\Delta$ 12 desaturase gene as described for the cDNA gene library. Autoradiography of the membranes showed 9 positive plaques. These plaques were isolated, titled and subjected to a second screening. 6 plaques were re-confirmed as positive. 4 of these gave a very strong signal.

#### Analysis of the clones isolated

The following positive phage clones were converted into plasmids by *in vivo* excision in accordance with the method suggested by Stratagene (Gigapack Gold in vitro packaging) : I, F, 4 (cDNA gene library), N2, N11, N17, N18, N21, N25 (genome gene library).

The plasmid DNA of the clones of the cDNA gene library was isolated and the length of the insert analyzed by digestion with Eco RI. The plasmid DNA of the genome clones was isolated, the length of the insert analyzed by cutting

with restriction enzyme, and the clones rechecked by hybridization with the *Arabidopsis* probe. Figure 1 shows the map of the N2 genome clone.

### Sequencing

The N2 clone was selected from the genome clones. For sequencing, the insert was fragmented with Sau3A restriction enzyme and the fragments obtained were subcloned in pUC18 vector cut with BamHI (Maniatis, Molecular cloning, 1989, Cold Spring Harbor Laboratory Press, 1.68-1.69). The clones obtained were analyzed both for the length of the insert and by hybridization with the *Arabidopsis* probe. Since the N2 insert was 2.8 kb and hence longer than the  $\Delta 12$  desaturase gene, the hybridization excluded the clones containing sequences outside the gene. The insert of the I, F, 4 and N2 clones was isolated and sequenced with the use of the Sequenase kit and (35S)dATP. All of the clones (cDNA and genome) were first sequenced at the ends with the use of primers which could couple with the vector in both orientations. In order to complete the internal regions and to assemble the fragments of the N2 genome clone, internal oligonucleotides were then designed and synthesized and were used for the sequencing. The following table shows the sequences of the internal oligonucleotides:

OLIGONUCLEOTIDE	SEQUENCE
N2-3SS	CAG ACC AGC ATC CGA GAC
N2-3SD	GGA TTG GCT TAG GGG GGC
N2-29R'S	GCC AAC CAT GTC ATC AAC CC
NOCCS	ATG GTA GAG AAG AGA TGG TG
COL	CTG GTG GGT TGT TGA AG
N2-S1N	GGA GAG GTC ATA AAC AAC

The I and F clones were sequenced entirely. As far as the N2 clone is concerned, only the regions corresponding to the gene were sequenced. Figures 2 and 3 show their sequence. The I and F cDNA clones were identical. A comparison between I and the N2 genome clone showed the same sequence (Fig. 4), indicating that N2 contains the gene which codes for the cDNA of the I clone.

### Comparison between the gene isolated and other desaturases

The nucleotide and amino-acid sequence of the N2 clone was compared with other desaturases (Figure 6). The greatest homology was with the two  $\Delta 12$  desaturases of the endoplasmic reticulum and with a hydroxylase of ricin which uses the same substrate as  $\Delta 12$  desaturase. Homology with the plastid  $\Delta 12$  desaturases and with both the plastid and endoplasmic reticulum  $\Delta 15$  desaturases was, however, much lower. Figure 5 shows the comparison between the amino-acid sequence of hazel  $\Delta 12$  and those of *Arabidopsis* and soya.

### Checking of the expression of the hazel $\Delta 12$ desaturase gene

RNA was extracted from kernels of the San Giovanni, Montebello and Nocchione varieties and from leaves of the Montebello and Nocchione varieties. After separation on agarose gel, the RNA was transferred onto a nylon membrane and hybridized with the insert of the I clone marked with digoxigenin. The result is shown in Figure 7, in which a band is visible in the kernel RNA but not in that of the leaves.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT:

(A) NAME: SOREMARTEC S.A.  
 (B) STREET: Dreve de l'Arc-en-Ciel 102  
 (C) CITY: Arlon-Schoppach  
 (E) COUNTRY: Belgium  
 (F) POSTAL CODE (ZIP): 6700

(ii) TITLE OF INVENTION: Isolation and sequencing of the  
 hazel FAD2-N  
 gene

(iii) NUMBER OF SEQUENCES: 4

(iv) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk  
 (B) COMPUTER: IBM PC compatible  
 (C) OPERATING SYSTEM: PC-DOS/MS-DOS  
 (D) SOFTWARE: PatentIn Release #1.0, Version #1.30

(EPO)

(vi) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: CH 0550/96  
 (B) FILING DATE: 04-MAR-1996

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1662 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Corylus avellana cv. Nocchione  
 (F) TISSUE TYPE: leaves

(vii) IMMEDIATE SOURCE:

(B) CLONE: N2

(ix) FEATURE:

(A) NAME/KEY: CDS  
 (B) LOCATION: 222..1370  
 (D) OTHER INFORMATION: /product= "delta-12 desaturase"  
 /gene= ""Fad2""

EP 0 794 250 A1

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

5 CCTCATAAAA AAGTAAGCTC ATTTACCTCA AGTAGGGTTT CCTTATGACA AATGAGTCCC  
60

GCAATCCTTT TCTATGAGGT GCTATAATTG CAAATGTCCA AATCATAGGG ATATGGATCC  
120

10 AAATACTATT AATATTATGT AGTGTGTTTT TTTTTTCCC TCAAATTTAC TCTCACACCT  
180

AAGTTGATTT TCTCCAGCAT TGGACATAGC CTCTGTAGAC A ATG GGA GCT AGA  
233

Met Gly Ala Arg  
1

15 AGC CGA ATG CCT GCT ACC AAC AAG CCT AAA GAG CAA AAA ACA CCC ATC  
281

Ser Arg Met Pro Ala Thr Asn Lys Pro Lys Glu Gln Lys Thr Pro Ile

20 5 10 15 20

CAG CGA GCA CCA CAC ACA AAA CCC CCA TTC ACT CTT AGC CAA CTC AAG  
329

Gln Arg Ala Pro His Thr Lys Pro Pro Phe Thr Leu Ser Gln Leu Lys

25 25 30 35

AAA GCC GTC CCA CCC AAT TGT TTC CAA CGC TCT CTC CTA CGC TCG TTC  
377

Lys Ala Val Pro Pro Asn Cys Phe Gln Arg Ser Leu Leu Arg Ser Phe

30 40 45 50

TCA TAT GTT GTT TAT GAC CTC TCC TTA GCC TTC CTC TTC TAC TAT ATT  
425

Ser Tyr Val Val Tyr Asp Leu Ser Leu Ala Phe Leu Phe Tyr Tyr Ile

35 55 60 65

GCT ACC TCT TAC TTC CAT CTC CTC CCT CAC CCC CTT TCC TAC TTG GCA  
473

Ala Thr Ser Tyr Phe His Leu Leu Pro His Pro Leu Ser Tyr Leu Ala

40 70 75 80

TGG TCA ATC TAT TGG GCT CTC CAA GGC TGC ATT CTC ACC GGC GTT TGG  
521

Trp Ser Ile Tyr Trp Ala Leu Gln Gly Cys Ile Leu Thr Gly Val Trp

50 85 90 95 100

55



EP 0 794 250 A1

5 GTC ATC GCA CAT GAG TGC GGT CAC CAT GCC TTT AGT GAC TAC CAA TGG  
 569  
 Val Ile Ala His Glu Cys Gly His His Ala Phe Ser Asp Tyr Gln Trp  
 105 110 115

10 GTT GAT GAC ATG GTT GGC CTA ACC CTT CAC TCT GCT CTT TTA GTT CCA  
 617  
 Val Asp Asp Met Val Gly Leu Thr Leu His Ser Ala Leu Leu Val Pro  
 120 125 130

15 TAC TTT TCA TGG AAG ATT AGC CAC TGT CGC CAC CAC TCT AAC ACC GGC  
 665  
 Tyr Phe Ser Trp Lys Ile Ser His Cys Arg His His Ser Asn Thr Gly  
 135 140 145

20 TCC CTT GAC CGA GAT GAG GTG TTT GTC CCC AAG CCG AAA TCC AAA ATG  
 713  
 Ser Leu Asp Arg Asp Glu Val Phe Val Pro Lys Pro Lys Ser Lys Met  
 150 155 160

25 CCA TGG TTT TCT AAG TAC TTC AAC AAC CCA CCA GGT AGG GTC CTC ACT  
 761  
 Pro Trp Phe Ser Lys Tyr Phe Asn Asn Pro Pro Gly Arg Val Leu Thr  
 165 170 175 180

30 CTT TTG ATC ACA CTC ACT CTA GGC TGG CCC TTG TAC TTA GCC TTG AAT  
 809  
 Leu Leu Ile Thr Leu Thr Leu Gly Trp Pro Leu Tyr Leu Ala Leu Asn  
 185 190 195

40 GTT TCT GGC CGA CCC TAT GAT CGT TTT GCT TGC CAC TAT GAT CCC TAT  
 857  
 Val Ser Gly Arg Pro Tyr Asp Arg Phe Ala Cys His Tyr Asp Pro Tyr  
 200 205 210

45 GGC CCC ATT TAT TCC AAT CGC GAA AGG TGT CAA ATA TTT GTC TCG GAT  
 905  
 Gly Pro Ile Tyr Ser Asn Arg Glu Arg Cys Gln Ile Phe Val Ser Asp  
 215 220 225

50 GCT GGT GTC TTT GCT ACA ACT TAT GTG CTT TAC TAC GCA GCA ATG TCA  
 953

55

EP 0 794 250 A1

Ala Gly Val Phe Ala Thr Thr Tyr Val Leu Tyr Tyr Ala Ala Met Ser

230

235

240

5

AAA GGG CTG GCA TGG CTT GTA TTC ATT TAT GGT ATG CCA TTG CTC ATA  
1001

Lys Gly Leu Ala Trp Leu Val Phe Ile Tyr Gly Met Pro Leu Leu Ile

10

245

250

255

260

GTG AAT GGC TTC CTT GTA TTA ATC ACC TAC TTG CAG CAC ACT CAC CCT  
1049

Val Asn Gly Phe Leu Val Leu Ile Thr Tyr Leu Gln His Thr His Pro

15

265

270

275

GCA TTG CCG CAC TAT GAC TCA TCA GAA TGG GAT TGG CTT AGG GGG GCA  
1097

Ala Leu Pro His Tyr Asp Ser Ser Glu Trp Asp Trp Leu Arg Gly Ala

20

280

285

290

TTG GCG ACG GCG GAT AGA GAT TAC GGA ATG CTG AAT AAG GTT TTC CAC  
1145

Leu Ala Thr Ala Asp Arg Asp Tyr Gly Met Leu Asn Lys Val Phe His

25

295

300

305

AAT ATC ATA GAC ACC CAT GTG GCT CAC CAT CTC TTC TCT ACC ATG CCT  
1193

Asn Ile Ile Asp Thr His Val Ala His His Leu Phe Ser Thr Met Pro

30

310

315

320

CAT TAC CAT GCA ATG GAA GCC ACC AAA GCA ATC AAG TCA ATA TTG GGC  
1241

His Tyr His Ala Met Glu Ala Thr Lys Ala Ile Lys Ser Ile Leu Gly

35

325

330

335

340

AAA TAC TAC CAG TTT GAT GGC ACT CCA GTT TAC AAG GCA GTG TGG AGG  
1289

Lys Tyr Tyr Gln Phe Asp Gly Thr Pro Val Tyr Lys Ala Val Trp Arg

40

345

350

355

GAG GCT AAA GAG TGC CTT TAT GTT GAG TCG GAC GAG GGG GCC CCT AAC  
1337

Glu Ala Lys Glu Cys Leu Tyr Val Glu Ser Asp Glu Gly Ala Pro Asn

45

50

EP 0 794 250 A1

360 365 370

5 AAA GGT GTT TTC TGG TAT CAG AGC AAG CTG TGA TATTGGCTGG ATAGAGCCAA  
1390  
Lys Gly Val Phe Trp Tyr Gln Ser Lys Leu \*

375 380

10 AGAAAATGTG ATTAGTAAGG TAGTGTCTTT GGTCAGTTTG GTGTGTTAAG GAACAAATAA  
1450

15 TAATAATTAG CGACTATGAA TAGTTATTGT TAAACAAAAT TCACCCTTAT GTTTAGCAGG  
1510

AACTTTTCTG GCTACACTTT TTTTCGTATG AAAAGCGCAT ATTTTTTAAT TGTTATATTG  
1570

20 TTTTGACATT ACTCAAGCTT CAAAATTAAT ATCACAGAAA ATATCCAATG TCGAAGGTTT  
1630

CATTGTAGGT TGAAAACCTT ATATTGAGGT GG  
1662

25 (2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 383 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

30 (ii) MOLECULE TYPE: protein  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

35 Met Gly Ala Arg Ser Arg Met Pro Ala Thr Asn Lys Pro Lys Glu Gln  
1 5 10 15

Lys Thr Pro Ile Gln Arg Ala Pro His Thr Lys Pro Pro Phe Thr Leu  
20 25 30

40 Ser Gln Leu Lys Lys Ala Val Pro Pro Asn Cys Phe Gln Arg Ser Leu  
35 40 45

45 Leu Arg Ser Phe Ser Tyr Val Val Tyr Asp Leu Ser Leu Ala Phe Leu  
50 55 60

50 Phe Tyr Tyr Ile Ala Thr Ser Tyr Phe His Leu Leu Pro His Pro Leu  
65 70 75 80

Ser Tyr Leu Ala Trp Ser Ile Tyr Trp Ala Leu Gln Gly Cys Ile Leu  
85 90 95

55

# EP 0 794 250 A1

Thr Gly Val Trp Val Ile Ala His Glu Cys Gly His His Ala Phe Ser  
 100 105 110  
 5 Asp Tyr Gln Trp Val Asp Asp Met Val Gly Leu Thr Leu His Ser Ala  
 115 120 125  
 10 Leu Leu Val Pro Tyr Phe Ser Trp Lys Ile Ser His Cys Arg His His  
 130 135 140  
 Ser Asn Thr Gly Ser Leu Asp Arg Asp Glu Val Phe Val Pro Lys Pro  
 145 150 155 160  
 15 Lys Ser Lys Met Pro Trp Phe Ser Lys Tyr Phe Asn Asn Pro Pro Gly  
 165 170 175  
 20 Arg Val Leu Thr Leu Leu Ile Thr Leu Thr Leu Gly Trp Pro Leu Tyr  
 180 185 190  
 Leu Ala Leu Asn Val Ser Gly Arg Pro Tyr Asp Arg Phe Ala Cys His  
 195 200 205  
 25 Tyr Asp Pro Tyr Gly Pro Ile Tyr Ser Asn Arg Glu Arg Cys Gln Ile  
 210 215 220  
 30 Phe Val Ser Asp Ala Gly Val Phe Ala Thr Thr Tyr Val Leu Tyr Tyr  
 225 230 235 240  
 Ala Ala Met Ser Lys Gly Leu Ala Trp Leu Val Phe Ile Tyr Gly Met  
 245 250 255  
 35 Pro Leu Leu Ile Val Asn Gly Phe Leu Val Leu Ile Thr Tyr Leu Gln  
 260 265 270  
 40 His Thr His Pro Ala Leu Pro His Tyr Asp Ser Ser Glu Trp Asp Trp  
 275 280 285  
 Leu Arg Gly Ala Leu Ala Thr Ala Asp Arg Asp Tyr Gly Met Leu Asn  
 290 295 300  
 45 Lys Val Phe His Asn Ile Ile Asp Thr His Val Ala His His Leu Phe  
 305 310 315 320  
 Ser Thr Met Pro His Tyr His Ala Met Glu Ala Thr Lys Ala Ile Lys  
 325 330 335  
 50 Ser Ile Leu Gly Lys Tyr Tyr Gln Phe Asp Gly Thr Pro Val Tyr Lys  
 55

340

345

350

5 Ala Val Trp Arg Glu Ala Lys Glu Cys Leu Tyr Val Glu Ser Asp Glu  
                   355                                  360                                  365

10 Gly Ala Pro Asn Lys Gly Val Phe Trp Tyr Gln Ser Lys Leu \*  
                   370                                  375                                  380

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1133 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(v) FRAGMENT TYPE: C-terminal

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Corylus avellana* L. cv. San Giovanni  
 (D) DEVELOPMENTAL STAGE: Seed, storage deposition stage

(vii) IMMEDIATE SOURCE:

- (B) CLONE: I

(ix) FEATURE:

- (A) NAME/KEY: mRNA  
 (B) LOCATION:1..1133  
 (D) OTHER INFORMATION:/partial  
                   /gene= "Fad2"

(ix) FEATURE:

- (A) NAME/KEY: CDS  
 (B) LOCATION:1..1019  
 (D) OTHER INFORMATION:/partial  
                   /codon\_start= 3  
                   /product= "delta-12 desaturase"  
                   /gene= "Fad2"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

45 TC CAA CGC TCT CTC CTA CGC TCG TTC TCA TAT GTT GTT TAT GAC CTC  
                   47

Gln Arg Ser Leu Leu Arg Ser Phe Ser Tyr Val Val Tyr Asp Leu

385

390

395

55 TCC TTA GCC TTC CTC TTC TAC TAT ATT GCT ACC TCT TAC TTC CAT CTC

EP 0 794 250 A1

95  
 Ser Leu Ala Phe Leu Phe Tyr Tyr Ile Ala Thr Ser Tyr Phe His Leu  
 400 405 410  
 5  
 CTC CCT CAC CCC CTT TCC TAC TTG GCA TGG TCA ATC TAT TGG GCT CTC  
 143  
 Leu Pro His Pro Leu Ser Tyr Leu Ala Trp Ser Ile Tyr Trp Ala Leu  
 415 420 425 430  
 10  
 CAA GGC TGC ATT CTC ACC GGC GTT TGG GTC ATC GCA CAT GAG TGC GGT  
 191  
 Gln Gly Cys Ile Leu Thr Gly Val Trp Val Ile Ala His Glu Cys Gly  
 435 440 445  
 15  
 CAC CAT GCC TTT AGT GAC TAC CAA TGG GTT GAT GAC ATG GTT GGC CTA  
 239  
 His His Ala Phe Ser Asp Tyr Gln Trp Val Asp Asp Met Val Gly Leu  
 450 455 460  
 20  
 ACC CTT CAC TCT GCT CTT TTA GTT CCA TAC TTT TCA TGG AAG ATT AGC  
 287  
 Thr Leu His Ser Ala Leu Leu Val Pro Tyr Phe Ser Trp Lys Ile Ser  
 465 470 475  
 25  
 CAC TGT CGC CAC CAC TCT AAC ACC GGC TCC CTT GAC CGA GAT GAG GTG  
 335  
 His Cys Arg His His Ser Asn Thr Gly Ser Leu Asp Arg Asp Glu Val  
 480 485 490  
 30  
 TTT GTC CCC AAG CCG AAA TCC AAA ATG CCA TGG TTT TCT AAG TAC TTC  
 383  
 Phe Val Pro Lys Pro Lys Ser Lys Met Pro Trp Phe Ser Lys Tyr Phe  
 495 500 505 510  
 35  
 AAC AAC CCA CCA GGT AGG GTC CTC ACT CTT TTG ATC ACA CTC ACT CTA  
 431  
 Asn Asn Pro Pro Gly Arg Val Leu Thr Leu Leu Ile Thr Leu Thr Leu  
 515 520 525  
 40  
 GGC TGG CCC TTG TAC TTA GCC TTG AAT GTT TCT GGC CGA CCC TAT GAT  
 479  
 Gly Trp Pro Leu Tyr Leu Ala Leu Asn Val Ser Gly Arg Pro Tyr Asp  
 55

EP 0 794 250 A1

	530	535	540
5	CGT TTT GCT TGC CAC TAT GAT CCC TAT GGC CCC ATT TAT TCC AAT CGC 527 Arg Phe Ala Cys His Tyr Asp Pro Tyr Gly Pro Ile Tyr Ser Asn Arg	545	550 555
10	GAA AGG TGT CAA ATA TTT GTC TCG GAT GCT GGT GTC TTT GCT ACA ACT 575 Glu Arg Cys Gln Ile Phe Val Ser Asp Ala Gly Val Phe Ala Thr Thr	560	565 570
15	TAT GTG CTT TAC TAC GCA GCA ATG TCA AAA GGG CTG GCA TGG CTT GTA 623 Tyr Val Leu Tyr Tyr Ala Ala Met Ser Lys Gly Leu Ala Trp Leu Val	575	580 585 590
20	TTC ATT TAT GGT ATG CCA TTG CTC ATA GTG AAT GGC TTC CTT GTA TTA 671 Phe Ile Tyr Gly Met Pro Leu Leu Ile Val Asn Gly Phe Leu Val Leu	595	600 605
25	ATC ACC TAC TTG CAG CAC ACT CAC CCT GCA TTG CCG CAC TAT GAC TCA 719 Ile Thr Tyr Leu Gln His Thr His Pro Ala Leu Pro His Tyr Asp Ser	610	615 620
30	TCA GAA TGG GAT TGG CTT AGG GGG GCA TTG GCG ACG GCG GAT AGA GAT 767 Ser Glu Trp Asp Trp Leu Arg Gly Ala Leu Ala Thr Ala Asp Arg Asp	625	630 635
35	TAC GGA ATG CTG AAT AAG GTT TTC CAC AAT ATC ATA GAC ACC CAT GTG 815 Tyr Gly Met Leu Asn Lys Val Phe His Asn Ile Ile Asp Thr His Val	640	645 650
40	GCT CAC CAT CTC TTC TCT ACC ATG CCT CAT TAC CAT GCA ATG GAA GCC 863 Ala His His Leu Phe Ser Thr Met Pro His Tyr His Ala Met Glu Ala	655	660 665 670
45			
50			
55			

# EP 0 794 250 A1

ACC AAA GCA ATC AAG TCA ATA TTG GGC AAA TAC TAC CAG TTT GAT GGC  
911  
Thr Lys Ala Ile Lys Ser Ile Leu Gly Lys Tyr Tyr Gln Phe Asp Gly

5 675 680 685

ACT CCA GTT TAC AAG GCA GTG TGG AGG GAG GCT AAA GAG TGC CTT TAT  
959  
Thr Pro Val Tyr Lys Ala Val Trp Arg Glu Ala Lys Glu Cys Leu Tyr

10 690 695 700

GTT GAG TCG GAC GAG GGG GCC CCT AAC AAA GGT GTT TTC TGG TAT CAG  
1007  
Val Glu Ser Asp Glu Gly Ala Pro Asn Lys Gly Val Phe Trp Tyr Gln

15 705 710 715

AGC AAG CTG TGA TATTGGCTGG ATAGAGCCAA AGAAAATGTG ATTAGTAAGG  
1059

Ser Lys Leu \*  
720

25 TAGTGTCTTT GGTCAGTTG GTGTGTTAAG GAACAAATAA TAATAATTAG CGACTATGAA  
1119

TAGTTATTGT TAAA  
1133

30

(2) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:

35 (A) LENGTH: 339 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

40 Gln Arg Ser Leu Leu Arg Ser Phe Ser Tyr Val Val Tyr Asp Leu Ser  
1 5 10 15

45 Leu Ala Phe Leu Phe Tyr Tyr Ile Ala Thr Ser Tyr Phe His Leu Leu  
20 25 30

Pro His Pro Leu Ser Tyr Leu Ala Trp Ser Ile Tyr Trp Ala Leu Gln  
35 40 45

50

Gly Cys Ile Leu Thr Gly Val Trp Val Ile Ala His Glu Cys Gly His

55



EP 0 794 250 A1

	50	55	60
5	His Ala Phe Ser Asp Tyr Gln Trp Val Asp Asp Met Val Gly Leu Thr 65 70 75 80		
	Leu His Ser Ala Leu Leu Val Pro Tyr Phe Ser Trp Lys Ile Ser His 85 90 95		
10	Cys Arg His His Ser Asn Thr Gly Ser Leu Asp Arg Asp Glu Val Phe 100 105 110		
15	Val Pro Lys Pro Lys Ser Lys Met Pro Trp Phe Ser Lys Tyr Phe Asn 115 120 125		
	Asn Pro Pro Gly Arg Val Leu Thr Leu Leu Ile Thr Leu Thr Leu Gly 130 135 140		
20	Trp Pro Leu Tyr Leu Ala Leu Asn Val Ser Gly Arg Pro Tyr Asp Arg 145 150 155 160		
25	Phe Ala Cys His Tyr Asp Pro Tyr Gly Pro Ile Tyr Ser Asn Arg Glu 165 170 175		
	Arg Cys Gln Ile Phe Val Ser Asp Ala Gly Val Phe Ala Thr Thr Tyr 180 185 190		
30	Val Leu Tyr Tyr Ala Ala Met Ser Lys Gly Leu Ala Trp Leu Val Phe 195 200 205		
35	Ile Tyr Gly Met Pro Leu Leu Ile Val Asn Gly Phe Leu Val Leu Ile 210 215 220		
	Thr Tyr Leu Gln His Thr His Pro Ala Leu Pro His Tyr Asp Ser Ser 225 230 235 240		
40	Glu Trp Asp Trp Leu Arg Gly Ala Leu Ala Thr Ala Asp Arg Asp Tyr 245 250 255		
45	Gly Met Leu Asn Lys Val Phe His Asn Ile Ile Asp Thr His Val Ala 260 265 270		
	His His Leu Phe Ser Thr Met Pro His Tyr His Ala Met Glu Ala Thr 275 280 285		
50	Lys Ala Ile Lys Ser Ile Leu Gly Lys Tyr Tyr Gln Phe Asp Gly Thr 290 295 300		
55			

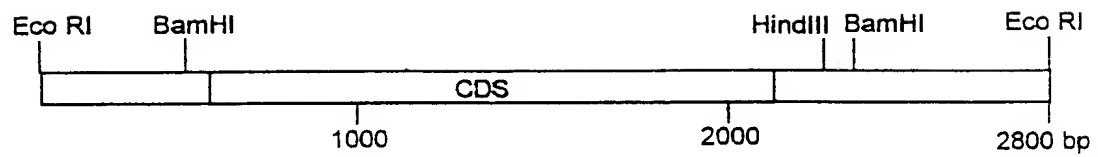
Pro Val Tyr Lys Ala Val Trp Arg Glu Ala Lys Glu Cys Leu Tyr Val  
 305 310 315 320

5 Glu Ser Asp Glu Gly Ala Pro Asn Lys Gly Val Phe Trp Tyr Gln Ser  
 325 330 335

10 Lys Leu \*

## 15 Claims

1. A fragment of DNA from hazel (*Corylus avellana* L.) comprising the nucleotide sequence shown in Figure 2.
2. A DNA fragment comprising the nucleotide sequence shown in Figure 2 from base 222 to base 1367, which codes  
 20 for the hazel  $\Delta$ 12 desaturase enzyme of the endoplasmic reticulum or for a homologous sequence which can code for the same amino-acid sequence.
3. A nucleotide sequence coding for a protein or peptide having an amino-acid homology greater than or equal to 80% and preferably greater than 90% with the hazel  $\Delta$ 12 desaturase enzyme of the endoplasmic reticulum of Claim 2  
 25 and having the function of the said enzyme.
4. A recombinant DNA sequence comprising a DNA sequence according to Claims 1, 2 and 3, or a portion of such a sequence, together with sequences regulating expression.
- 30 5. A recombinant DNA molecule comprising a cloning vector in which a DNA sequence according to any one of Claims 1, 2, 3 and 4 is inserted.
6. A DNA molecule according to Claim 5, in which the cloning vector is a plasmid or a phage.
- 35 7. A DNA molecule according to Claim 4 or Claim 5 having the restriction map shown in Figure 1.
8. A host organism including a recombinant DNA molecule according to any one of Claims 3 to 6.
9. A host organism according to Claim 8, selected from a vegetable cell, an animal cell, and a micro-organism.  
 40
10. A genetically modified organism capable of expressing the FAD2-N gene, having the amino-acid sequence shown in Figure 2 from bp 222 to bp 1367, portions of this gene, or this gene conjugated with other molecules and containing sequences which can inactivate endogenous genes.
- 45 11. A hazel  $\Delta$ 12 desaturase enzyme of the endoplasmic reticulum having the amino-acid sequence shown in Figure 2 in substantially pure form.
12. A fusion polypeptide comprising the amino-acid sequence of the enzyme of Claim 11, in which the amino-acids additively connected thereto do not interfere with the desaturase activity or can easily be eliminated.
- 50 13. The use of the FAD2-N gene coding for the hazel  $\Delta$ 12 desaturase enzyme of the endoplasmic reticulum or of portions thereof for the isolation of enzymes having the function of hazel desaturase or of the desaturase of another species.
- 55 14. The use of the nucleotide sequences of the FAD2-N gene shown in Figure 2 for the construction of expression systems which can alter the fatty-acid content in hazel.



**Fig. 1** - Restriction map of the genomic clone "N2". CDS: coding region; bp: base pair.

**Fig. 2** - Nucleotide sequence of the gene FAD2-N corresponding to an internal fragment of the genomic clone "N2". Aminoacid residues of the coding region are also reported.

```

CCTCATAAAAAAGTAAGCTCATTTACCTCAAGTAGGGTTTCCTTATGACAAATGAGTCCC 60
GGAGTATTTTTTCATTTCGAGTAAATGGAGTTCATCCCAAAGGAATACTGTTTACTCAGGG

GCAATCCTTTTCTATGAGGTGCTATAATTGCAAATGTCCAAATCATAGGGATATGGATCC 120
CGTTAGGAAAAGATACTCCACGATATTAACGTTTACAGGTTTAGTATCCCTATACCTAGG

AAATACTATTAATATTATGTAGTGTGTTTTTTTTTTTCCCTCAAATTTACTCTCACACCT 180
TTTATGATAATTATAATACATCACACAAAAAAGGGAGTTTAAATGAGAGTGTGGA

AAGTTGATTTTCTCCAGCATTGGACATAGCCTCTGTAGACAATGGGAGCTAGAAGCCGAA 240
TTCAACTAAAAGAGGTCGTAACCTGTATCGGAGACATCTGTTACCCTCGATCTTCGGCTT
Met Gly Ala Arg Ser Arg

TGCCTGCTACCAACAAGCCTAAAGAGCAAAAAACACCCATCCAGCGAGCACCACACACAA 300
ACGGACGATGGTTGTTTCGGATTTCTCGTTTTTTTGTGGGTAGGTCGCTCGTGGTGTGTGTT
Met Pro Ala Thr Asn Lys Pro Lys Glu Gln Lys Thr Pro Ile Gln Arg Ala Pro His Thr

AAGCCCCCATTCACTCTTAGCCAACTCAAGAAAGCCGTCCCACCCCAATTGTTTCCAACGCT 360
TTGGGGGTAAAGTGAGAATCGGTTGAGTTCTTTCGGCAGGGTGGGTAAACAAAGGTTGCGA
Lys Pro Pro Phe Thr Leu Ser Gln Leu Lys Lys Ala Val Pro Pro Asn Cys Phe Gln Arg

CTCTCCTACGCTCGTTCTCATATGTTGTTTATGACCTCTCCTTAGCCTTCCTCTTCTACT 420
GAGAGGATGCGAGCAAGAGTATACAACAAATACTGGAGAGGAATCGGAAGGAGAAGATGA
Ser Leu Leu Arg Ser Phe Ser Tyr Val Val Tyr Asp Leu Ser Leu Ala Phe Leu Phe Tyr

ATATTGCTACCTCTTACTTCCATCTCCTCCCTCACCCCTTTCTTACTTGGCATGGTCAA 480
TATAACGATGGAGAATGAAGGTAGAGGAGGGAGTGGGGGAAAGGATGAACCGTACCAGTT
Tyr Ile Ala Thr Ser Tyr Phe His Leu Leu Pro His Pro Leu Ser Tyr Leu Ala Trp Ser

TCTATTGGGCTCTCCAAGGCTGCATTCTCACCGGCGTTTGGGTTCATCGCACATGAGTGCG 540
AGATAACCCGAGAGGTTCCGACGTAAGAGTGGCCGCAAACCCAGTAGCGTGTACTCACGC
Ile Tyr Trp Ala Leu Gln Gly Cys Ile Leu Thr Gly Val Trp Val Ile Ala His Glu Cys

GTCACCATGCCTTTAGTGACTACCAATGGGTTGATGACATGGTTGGCCTAACCCCTTCACT 600
CAGTGGTACGGAAATCACTGATGGTTACCCAACTACTGTACCAACCGGATTGGGAAGTGA
Gly His His Ala Phe Ser Asp Tyr Gln Trp Val Asp Asp Met Val Gly Leu Thr Leu His

CTGCTCTTTTGTAGTTCCATACTTTTCATGGAAGATTAGCCACTGTGCGCCACCACTCTAACA 660
GACGAGAAAATCAAGGTATGAAAAGTACCTTCTAATCGGTGACAGCGGTGGTGAGATTGT
Ser Ala Leu Leu Val Pro Tyr Phe Ser Trp Lys Ile Ser His Cys Arg His His Ser Asn

```

CCGGCTCCCTTGACCGAGATGAGGTGTTTGTCCCCAAGCCGAAATCCAAAATGCCATGGT 720  
GGCCGAGGGAACCTGGCTCTACTCCACAAACAGGGGTTCGGCTTTAGGTTTTACGGTACCA

Thr Gly Ser Leu Asp Arg Asp Glu Val Phe Val Pro Lys Pro Lys Ser Lys Met Pro Trp

TTTCTAAGTACTTCAACAACCCACCAGGTAGGGTCCTCACTCTTTTGATCACACTCACTC 780  
AAAGATTCATGAAGTTGTTGGGTGGTCCATCCCAGGAGTGAGAAAACCTAGTGTGAGTGAG

Phe Ser Lys Tyr Phe Asn Asn Pro Pro Gly Arg Val Leu Thr Leu Leu Ile Thr Leu Thr

TAGGCTGGCCCTTGTACTTAGCCTTGAATGTTTCTGGCCGACCCTATGATCGTTTTGCTT 840  
ATCCGACCGGGAACATGAATCGGAACCTACAAAGACCGGCTGGGATACTAGCAAAACGAA

Leu Gly Trp Pro Leu Tyr Leu Ala Leu Asn Val Ser Gly Arg Pro Tyr Asp Arg Phe Ala

GCCACTATGATCCCTATGGCCCCATTTATTCCAATCGCGAAAGGTGTCAAATATTTGTCT 900  
CGGTGATACTAGGGATACCGGGGTAAATAAGGTTAGCGCTTCCACAGTTTATAAACAGA

Cys His Tyr Asp Pro Tyr Gly Pro Ile Tyr Ser Asn Arg Glu Arg Cys Gln Ile Phe Val

CGGATGCTGGTGTCTTTGCTACAACCTTATGTGCTTTACTACGCAGCAATGTCAAAGGGC 960  
GCCTACGACCACAGAAACGATGTTGAATACAGAAATGATGCGTCGTTACAGTTTTCCCG

Ser Asp Ala Gly Val Phe Ala Thr Thr Tyr Val Leu Tyr Tyr Ala Ala Met Ser Lys Gly

TGGCATGGCTTGTATTCATTTATGGTATGCCATTGCTCATAGTGAATGGCTTCCTTGTAT 1020  
ACCGTACCGAACATAAGTAAATACCATACGGTAACGAGTATCACTTACCGAAGGAACATA

Leu Ala Trp Leu Val Phe Ile Tyr Gly Met Pro Leu Leu Ile Val Asn Gly Phe Leu Val

TAATCACCTACTTGCAGCACACTCACCCCTGCATTGCCGCACTATGACTCATCAGAATGGG 1080  
ATTAGTGGATGAACGTCGTGTGAGTGGGACGTAACGGCGTGATACTGAGTAGTCTTACCC

Leu Ile Thr Tyr Leu Gln His Thr His Pro Ala Leu Pro His Tyr Asp Ser Ser Glu Trp

ATTGGCTTAGGGGGGCATTGGCGACGGCGGATAGAGATTACGGAATGCTGAATAAGGTTT 1140  
TAACCGAATCCCCCGTAACCGCTGCCGCCTATCTCTAATGCCTTACGACTTATTCCAA

Asp Trp Leu Arg Gly Ala Leu Ala Thr Ala Asp Arg Asp Tyr Gly Met Leu Asn Lys Val

TCCACAATATCATAGACACCCATGTGGCTCACCATCTCTTCTCTACCATGCCTCATTACC 1200  
AGGTGTTATAGTATCTGTGGGTACACCGAGTGGTAGAGAAGAGATGGTACGGAGTAATGG

Phe His Asn Ile Ile Asp Thr His Val Ala His His Leu Phe Ser Thr Met Pro His Tyr

ATGCAATGGAAGCCACCAAAGCAATCAAGTCAATATTGGGCAAATACTACCAGTTTGATG 1260  
TACGTTACCTTCGGTGGTTTCGTTAGTTTCAAGTTATAACCCGTTTATGATGGTCAAACCTAC

His Ala Met Glu Ala Thr Lys Ala Ile Lys Ser Ile Leu Gly Lys Tyr Tyr Gln Phe Asp

GCACTCCAGTTTACAAGGCAGTGTGGAGGGAGGCTAAAGAGTGCCTTTATGTTGAGTCGG 1320  
CGTGAGGTCAAATGTTCCGTCACACCTCCCTCCGATTTCTCACGGAAATACAACCTCAGCC

Gly Thr Pro Val Tyr Lys Ala Val Trp Arg Glu Ala Lys Glu Cys Leu Tyr Val Glu Ser

ACGAGGGGGCCCCCTAACAAAGGTGTTTTCTGGTATCAGAGCAAGCTGTGATATTGGCTGG 138C  
 TGCTCCCCCGGGGATTGTTTCCACAAAAGACCATAGTCTCGTTTCGACACTATAACCGACC  
 Asp Glu Gly Ala Pro Asn Lys Gly Val Phe Trp Tyr Gln Ser Lys Leu

ATAGAGCCAAAGAAAATGTGATTAGTAAGGTAGTGTCTTTGGTCAGTTTGGTGTGTTAAG 144C  
 TATCTCGGTTTCTTTTACACTAATCATTCCATCACAGAAACCAGTCAAACCACACAATTC

GAACAAATAATAATAATTAGCGACTATGAATAGTTATTGTTAAACAAAATTCACCCCTTAT 150C  
 CTTGTTTATTATTATTAATCGCTGATACTTATCAATAACAATTTGTTTTAAGTGGGAATA

GTTTAGCAGGAACTTTTCTGGCTACACTTTTTTTTCGTATGAAAAGCGCATATTTTTTTAAT 156C  
 CAAATCGTCCTTGAAAAGACCGATGTGAAAAAAGCATACTTTTCGCGTATAAAAAATTA

TGTTATATTGTTTTGACATTACTCAAGCTTCAAAATTAATATCACAGAAAATATCCAATG 162C  
 ACAATATAACAAAACGTAAATGAGTTCGAAGTTTTAATTATAGTGTCTTTTATAGGTTAC

TCGAAGGTTTCATTGTAGGTTGAAAACTTTATATTGAGGTGG 168C  
 AGCTTCCAAAGTAACATCCAACTTTTGAAATATAACTCCACC

10 20 30 40 50  
 tccaaagctctctctacgctcggttctcatatggtgtttatgacctctcc 50  
 tttagccttccctcttctactatattgctacctcttacctccatctccccc 100  
 tcccccccttctctacttggcatggtcacatctattgggctctccaaaggt 150  
 gcattctcaccggcggttgggtcctcgccacatgagtgaggcaccatgcc 200  
 tttagtgaactacacatgggttgatgacatgggtgggctcaccctctccctc 250  
 tgcctctttagtccatacttttccatgggaagattagccactgtcgccccc 300  
 accttaaacaccggctccccttgaccgagatgagggtgtttgtccccacgcg 350  
 aactccaaacatgccatgggtttcttaagtaacttcaaacacccacaggtag 400  
 ggctctccctcttttgcacacacacacacacacacacacacacacacacac 450  
 ccttgcaatgttcttgcccgacccctatgatgttttgcttgccacacatgat 500  
 ccttatggccccattttattccaatcgcgaaagggtgtcaaatatttgctctc 550  
 ggatgctgggtgtctttgtctccaaacttatgtgtctttactacgcagcaatgt 600  
 caaaaggggtggcatgggttgctattcatttatgggtatgcaattgctctctc 650  
 gtgaatgggttccctgttatataacacacacacacacacacacacacacac 700  
 attgcacgac 750  
 cgaagggcggtatgagattacgggaatgctgaatcagggtttccacacacacac 800  
 atagac 850  
 tgcaatgggaagccccccccacacacacacacacacacacacacacacacacac 900  
 agtttgcctggcactccagtttccaaagggcagtggtggggggggggcaccagag 950  
 tgcaatttatgttgagtcgggaaggggggggggggggggggggggggggggg 1000  
 gtatcagagccagctgtgtgatattgggtgggtatgggggggggggggggggg 1050  
 ttatgtaaggtatgtgtctttgggtcaggtttgggtgtgttccgggggggggggg 1100  
 aatcaattgcgacatgatgcaatgattatgtttcc 1133

Fig. 3 - Nucleotide sequence of cDNA clone "I".

**Fig 4** - Nucleotide sequence alignment of clones "I" (I.SEQ) and "N2" (N2.SEQ).

```

1  ----- I.SEQ
1  CCTCATAAAAAAGTAAGCTCATTTACCTCAAGTAGGGTTT N2.SEQ

1  ----- I.SEQ
41 CCTTATGACAAATGAGTCCCGCAATCCTTTTCTATGAGGT N2.SEQ

1  ----- I.SEQ
81 GCTATAATTGCAAATGTCCAAATCATAGGGATATGGATCC N2.SEQ

1  ----- I.SEQ
121 AAATACTATTATAATTATGTAGTGTGTTTTTTTTTCCC N2.SEQ

1  ----- I.SEQ
161 TCAAATTTACTCTCACACCTAAGTTGATTTTCTCCAGCAT N2.SEQ

1  ----- I.SEQ
201 TGGACATAGCCTCTGTAGACAAATGGGAGCTAGAAGCCGAA N2.SEQ

1  ----- I.SEQ
241 TGCTTGCTACCAACAAGCCTAAAGAGCAAAAAACACCCAT N2.SEQ

1  ----- I.SEQ
281 CCAGCGAGCACCCACACACAAAACCCCATTCACCTCTTAGC N2.SEQ

1  ----- TCCAAACGCT I.SEQ
321 CAACTCAAGAAAGCCGCTCCCAACCAATTGTTTCCAAACGCT N2.SEQ

10  CTCTCCTACGCTCGTTCTCATATGTTGTTTATGACCTCTC I.SEQ
361 CTCTCCTACGCTCGTTCTCATATGTTGTTTATGACCTCTC N2.SEQ

50  CTTAGCCTTCCCTCTTCTACTATATTGCTACCTCTTACTTTC I.SEQ
401 CTTAGCCTTCCCTCTTCTACTATATTGCTACCTCTTACTTTC N2.SEQ

90  CATCTCCTCCCTCACCCCTTTTCTACTTGGCATGGTCAA I.SEQ
441 CATCTCCTCCCTCACCCCTTTTCTACTTGGCATGGTCAA N2.SEQ

130 TCTATTGGGGCTCTCCAAGGCTGCATTCTCACCGGCGTTTG I.SEQ
481 TCTATTGGGGCTCTCCAAGGCTGCATTCTCACCGGCGTTTG N2.SEQ

170 GGTTCATCGCACATGAGTGCGGTACCATGCTTTTAGTGAC I.SEQ
521 GGTTCATCGCACATGAGTGCGGTACCATGCTTTTAGTGAC N2.SEQ

210 TACCAATGGGTTGATGACATGGTTGGCCTAACCCCTTCACT I.SEQ
561 TACCAATGGGTTGATGACATGGTTGGCCTAACCCCTTCACT N2.SEQ

250 CTGCTCTTTTAGTTCCATACTTTTTCATGGAAGATTAGCCA I.SEQ
601 CTGCTCTTTTAGTTCCATACTTTTTCATGGAAGATTAGCCA N2.SEQ

290 CTGTCCCCACCACTCTAACACCGGCTCCCTTGACCGAGAT I.SEQ
641 CTGTCCCCACCACTCTAACACCGGCTCCCTTGACCGAGAT N2.SEQ

330 GAGGTGTTTGTCCCCAAGCCGAAATCCAAAATGCCATGGT I.SEQ
681 GAGGTGTTTGTCCCCAAGCCGAAATCCAAAATGCCATGGT N2.SEQ

370 TTTCTAAGTACTTCAACAACCCACCGGTAGGGTCCTCAC I.SEQ
721 TTTCTAAGTACTTCAACAACCCACCGGTAGGGTCCTCAC N2.SEQ

410 TCTTTTGTATCACACTCACTCTAGGCTGGCCCTTGTACTTA I.SEQ
761 TCTTTTGTATCACACTCACTCTAGGCTGGCCCTTGTACTTA N2.SEQ

450 GCCTTGAAATGTTTCTGGCCGACCCATGATCGTTTTGCTT I.SEQ
801 GCCTTGAAATGTTTCTGGCCGACCCATGATCGTTTTGCTT N2.SEQ

```



EP 0 794 250 A1

490 G C C A C T A T G A T C C C T A T G G C C C C A T T T A T T C C A A T C G C G A I.SEQ  
841 G C C A C T A T G A T C C C T A T G G C C C C A T T T A T T C C A A T C G C G A N2.SEQ

530 A A G G T G T C A A A T A T T T G T C T C G G A T G C T G G T G T C T T T G C T I.SEQ  
881 A A G G T G T C A A A T A T T T G T C T C G G A T G C T G G T G T C T T T G C T N2.SEQ

570 A C A A C T T A T G T G C T T T A C T A C G C A G C A A T G T C A A A A G G G C I.SEQ  
921 A C A A C T T A T G T G C T T T A C T A C G C A G C A A T G T C A A A A G G G C N2.SEQ

610 T G G C A T G G C T T G T A T T C A T T T A T G G T A T G C C A T T G C T C A T I.SEQ  
961 T G G C A T G G C T T G T A T T C A T T T A T G G T A T G C C A T T G C T C A T N2.SEQ

650 A G T G A A T G G C T T C C T T G T A T T A A T C A C C T A C T T G C A G C A C I.SEQ  
1001 A G T G A A T G G C T T C C T T G T A T T A A T C A C C T A C T T G C A G C A C N2.SEQ

690 A C T C A C C C T G C A T T G C C G C A C T A T G A C T C A T C A G A A T G G G I.SEQ  
1041 A C T C A C C C T G C A T T G C C G C A C T A T G A C T C A T C A G A A T G G G N2.SEQ

730 A T T G G C T T A G G G G G G C A T T G G C G A C G G C G G A T A G A G A T T A I.SEQ  
1081 A T T G G C T T A G G G G G G C A T T G G C G A C G G C G G A T A G A G A T T A N2.SEQ

770 C G G A A T G C T G A A T A A G G T T T T C C A C A A T A T C A T A G A C A C C I.SEQ  
1121 C G G A A T G C T G A A T A A G G T T T T C C A C A A T A T C A T A G A C A C C N2.SEQ

810 C A T G T G G C T C A C C A T C T C T T C T C T A C C A T G C C T C A T T A C C I.SEQ  
1161 C A T G T G G C T C A C C A T C T C T T C T C T A C C A T G C C T C A T T A C C N2.SEQ

850 A T G C A A T G G A A G C C A C C A A A G C A A T C A A G T C A A T A T T G G G I.SEQ  
1201 A T G C A A T G G A A G C C A C C A A A G C A A T C A A G T C A A T A T T G G G N2.SEQ

890 C A A A T A C T A C C A G T T T G A T G G C A C T C C A G T T T A C A A G G C A I.SEQ  
1241 C A A A T A C T A C C A G T T T G A T G G C A C T C C A G T T T A C A A G G C A N2.SEQ

930 G T G T G G A G G G A G G C T A A A G A G T G C C T T T A T G T T G A G T C G G I.SEQ  
1281 G T G T G G A G G G A G G C T A A A G A G T G C C T T T A T G T T G A G T C G G N2.SEQ

970 A C G A G G G G G C C C C T A A C A A A G G T G T T T T C T G G T A T C A G A G I.SEQ  
1321 A C G A G G G G G C C C C T A A C A A A G G T G T T T T C T G G T A T C A G A G N2.SEQ

1010 C A A G C T G T G A T A T T G G C T G G A T A G A G C C A A A G A A A A T G T G I.SEQ  
1361 C A A G C T G T G A T A T T G G C T G G A T A G A G C C A A A G A A A A T G T G N2.SEQ

1050 A T T A G T A A G G T A G T G T C T T T G G T C A G T T T G G T G T G T T A A G I.SEQ  
1401 A T T A G T A A G G T A G T G T C T T T G G T C A G T T T G G T G T G T T A A G N2.SEQ

1090 G A A C A A A T A A T A A T A A T T A G C G A C T A T G A A T A G T T A T T G T I.SEQ  
1441 G A A C A A A T A A T A A T A A T T A G C G A C T A T G A A T A G T T A T T G T N2.SEQ

1130 T A A A I.SEQ  
1481 T A A C A A A A T T C A C C C T T A T G T T T A G C A G G A A C T T T T C T G N2.SEQ

1133 I.SEQ  
1521 G C T A C A C T T T T T T T C G T A T G A A A A G C G C A T A T T T T T A A T N2.SEQ

1133 I.SEQ  
1561 T G T T A T A T T G T T T T G A C A T T A C T C A A G C T T C A A A A T T A A T N2.SEQ

1133 I.SEQ  
1601 A T C A C A G A A A A T A T C C A A T G T C G A A G G T T T C A T T G T A G G T N2.SEQ

1133 I.SEQ  
1641 T G A A A A C T T T A T A T T G A G G T G G N2.SEQ

```

1  M G A R S R M P I - A T N K P K E Q K T P I Q R A P H T K P P F T L S Q L K K A V N2.PRO
1  M G A G G R T D I V P I P A N R K S E V D P L K R V P F E K P Q F S L S Q I K K A I L43921.PRO
1  M G A G G R M P V P T S S K K S E I T D I T T K R V P C E K P P F S V G D L K K A I L26296.PRO

40 P P N C F Q R S L L R S F S Y V V Y D L S L A F L F Y Y I A T S Y F H L L P H P N2.PRO
41 P P H C F Q R S V L R S F S Y V V Y D L T I A F C L Y Y V A T H Y F H L L P G P L43921.PRO
41 P P H C F K R S I P R S F S Y L I S D I I I A S C F Y Y V A T N Y F S L L P Q P L26296.PRO

80 L S Y L A W S I Y W A L Q G C I L T G V W V I A H E C G H H A F S D Y Q W V D D N2.PRO
81 L S F R G M A I Y W A V Q G C I L T G V W V I A H E C G H H A F S D Y Q L L D D L43921.PRO
81 L S Y L A W P L Y W A C Q G C V L T G I W V I A H E C G H H A F S D Y Q W L D D L26296.PRO

120 M V G L T L H S A L L V P Y F S W K I S H C R H H S N T G S L D R D E V F V P K N2.PRO
121 I V G L I L H S A L L V P Y F S W K Y S H R R H H S N T G S L E R D E V F V P K L43921.PRO
121 T V G L I F H S F L L V P Y F S W K Y S H R R H H S N T G S L E R D E V F V P K L26296.PRO

160 P K S K M P W F S K Y F N N P P G R V L T L L I T L T L G W P L Y L A L N V S G N2.PRO
161 Q K S C I K W Y S K Y L N N P P G R V L T L A V T L T L G W P L Y L A L N V S G L43921.PRO
161 Q K S A I K W Y G K Y L N N P L G R I M M L T V Q F V L G W P L Y L A F N V S G L26296.PRO

200 R P Y D R F A C H Y D P Y G P I Y S N R E R C Q I F V S D A G V F A T T Y V L Y N2.PRO
201 R P Y D R F A C H Y D P Y G P I Y S D R E R L Q I Y I S D A G V L A V V Y G L F L43921.PRO
201 R P Y D G F A C H F F P N A P I Y N D R E R L Q I Y L S D A G I L A V C F G L Y L26296.PRO

240 Y A A M S K G L A W L V F I Y G M P L L I V N G F L V L I T Y L Q H T H P A L P N2.PRO
241 R L A M A K G L A W V V C V Y G V P L L V V N G F L V L I T F L Q H T H P A L P L43921.PRO
241 R Y A A A Q G M A S M I C L Y G V P L L I V N A F L V L I T Y L Q H T H P S L P L26296.PRO

280 H Y D S S E W D W L R G A L A T A D R D Y G M L N K V F H N I I D T H V A H H L N2.PRO
281 H Y T S S E W D W L R G A L A T V D R D Y G I L N K V F H N I T D T H V A H H L L43921.PRO
281 H Y D S S E W D W L R G A L A T V D R D Y G I L N K V F H N I T D T H V A H H L L26296.PRO

320 F S T M P H Y H A M E A T K A I K S I L G K Y Y Q F D G T P V Y K A V W R E A K N2.PRO
321 F S T M P H Y H A M E A T K A I K P I L G E Y Y R F D E T P F V K A M W R E A R L43921.PRO
321 F S T M P H Y N A M E A T K A I K P I L G D Y Y Q F D G T P W Y V A M Y R E A K L26296.PRO

360 E C L Y V E S D E G A P N K G V F W Y Q S K L N2.PRO
361 E C I Y V E P D Q S T E S K G V F W Y N N K L L43921.PRO
361 E C I Y V E P D R E G D K K G V Y W Y N N K L L26296.PRO

```

Fig. 5 - Aminoacid sequence alignment of  $\Delta 12$  desaturase from hazelnut (N2.PRO), *Arabidopsis* (L26296.PRO) and soybean (L43921.PRO). Homologous residues are boxed.

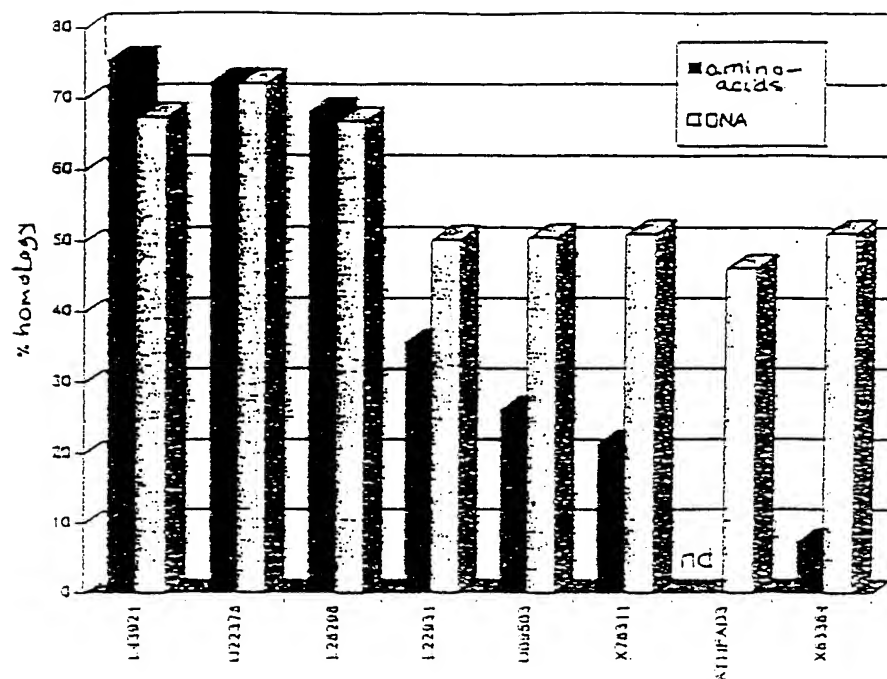


Fig. 6 - Homology between hazel  $\Delta 12$  desaturase and other desaturases

L43921:  $\Delta 12$  desaturase of the endoplasmic reticulum of soya  
 U22378:  $\Delta 12$  hydroxylase of ricin  
 L26296:  $\Delta 12$  desaturase of the endoplasmic reticulum of *Arabidopsis thaliana*  
 L22931:  $\Delta 15$  plastid desaturase of *Arabidopsis thaliana*  
 U09503:  $\Delta 12$  plastid desaturase of *Arabidopsis thaliana*  
 X78311:  $\Delta 12$  plastid desaturase of spinach  
 ATHFAD3:  $\Delta 15$  desaturase of the endoplasmic reticulum of *Arabidopsis thaliana*  
 X63364:  $\Delta 9$  plastid desaturase of rape

Note: nd: not determined since the amino-acid sequence is not known.

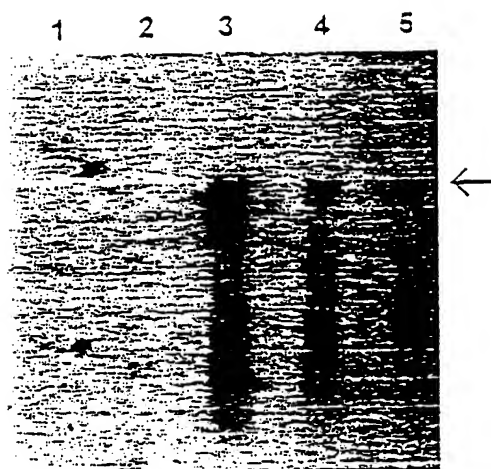


Fig. 7 - Northern blot of RNA of Montebello leaves (line 1), Nocchione leaves (line 2), Montebello kernels (line 3), Nocchione kernels (line 4), and San Giovanni kernels (line 5). The RNA was hybridized with the I clone of cDNA.



European Patent  
Office

# EUROPEAN SEARCH REPORT

Application Number  
EP 97 10 3098

DOCUMENTS CONSIDERED TO BE RELEVANT			
Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (Int.Cl.6)
X	WO 94 11516 A (DU PONT ;LIGHTNER JONATHAN EDWARD (US); OKULEY JOHN JOSEPH (US)) 26 May 1994 examples 1,6,7 ---	10,13	C12N15/53 C12N15/82 C12N9/02 C12N5/10 C12Q1/68 //A01H5/00
A,D	THE PLANT CELL, vol. 6, January 1994, pages 147-158, XP002034147 OKULEY, J., ET AL . : "ARABIDOPSIS FAD2 GENE ENCODES THE ENZYME THAT IS ESSENTIAL FOR POLYSATURATED LIPID SYNTHESIS" * page 155, column 2, line 28 * ---	1-14	
A	WO 95 22598 A (DU PONT ;LETO KENNETH JOSEPH (US); ULRICH JAMES FRANCIS (US)) 24 August 1995 * page 10, line 1 * -----	1-23	
			TECHNICAL FIELDS SEARCHED (Int.Cl.6)
			C12N
The present search report has been drawn up for all claims			
Place of search THE HAGUE		Date of completion of the search 3 July 1997	Examiner Holtorf, S
<p><b>CATEGORY OF CITED DOCUMENTS</b></p> <p>X : particularly relevant if taken alone Y : particularly relevant if combined with another document of the same category A : technological background O : non-written disclosure P : intermediate document</p> <p>T : theory or principle underlying the invention E : earlier patent document, but published on, or after the filing date D : document cited in the application L : document cited for other reasons ..... &amp; : member of the same patent family, corresponding document</p>			

EPO FORM 1503 03.92 (P04C01)